

## Materialize + dbt: Streaming for the modern data stack

Jessica Laughlin Materialize

## It's the modern data stack's world, we just live in it.



First, some definitions

### data stack → collection of all your data tools

### First, some definitions

data stack	$\rightarrow$	collection of all your data tools
modern data stack	$\rightarrow$	collection of all your data tools, <mark>in the</mark> cloud

What's so good about the modern data stack?

- More powerful
- Less expensive

What's so good about the modern data stack?

- More powerful
- Less expensive
- SQL interface

## A small detour: interfaces are important.

SAME AND	2.10	57	ACTIVITY OF A	152.28	di No	1	816	5 (A)	11	5.3.4	140	di l	10.0	A-C					-	-	-					-	44	SU.	876	160	076	N-84
	100	1996) 1996)	Reservation Managements	Contraction of the	-		-																									
1920 ALEXAND																				0	OM	MU	NE	D		-	1	~		-		5
23.350 4																						-		-	~	-	1	au		er	- e	/
402626																		-				-			-	E	I	N	9	E	25	3
	-	-		1	1			-	EX	TRETI	13				-			100	dho	NES I	REP.	AILAT	IONS	•			T	-	-			
			•	1000	-	1	INTE	PARAE PO	A FEER	Ja	mer_		TRAXA	15.25	REGRE	E	-	tabels of	-					TRAVA	13.1	es mon		Inter	Pine I	-		
	CLASS	Chatest	INDRCATION N	BATES	Cedies	ulers _		Crida .	4 1	1-11-15	639	Jr.		ader		-		I men a	1	-	TUT	AL.			an	-			Center	64-		TOT
1000 4104	08.95	A \$600	DES DÉPENSES	den .	-		A150X	TRATA	13	Rahai	in ride		le6	39.4	1e	TRAT	ALS.	TEAS BO	NEX	-	data a	evan	-	-	-			TRAVET	71	aval t		Falses
SPON DA	NING	SUM	et des pièces admissère	PECIS	de	1 te	abain abain drdoi	termin raba	da s	Neterr	Are	ret	Nature		Argent	east of	ale Miket	rah ros é	in Isat	Sale	-	Arge	ret	Note	100	Arpst	1	Poliets Bros dide	1 20	aluin a didei	1 3	Non
San Harris			and the second		Spot	-	-					_	1		-	-	-	-	1	200			-	8	-	-		11		-		_
Sal and a set																																
			Esport des années antérieures														_			-		_										
																											1					
			Infinis de Terminia	mini	90															1											4	
	17	2	Binnes & Drangen &	s.					-																						.4	
199999	25	3	Inlana - de fission	capionia	90											-				•		-	•								-	
	-	4	Arlani de Star	3s.Marss	.90																							-		-		
			5		0.72		+				100						1						12							1	T	
	-	c	Coliner on 31 Mars	and	15		1	1			1.																					
	85	6	Destations	3450									4	50											1.							
	43	7	Sulan a stra	e	45											-																
	14	1	Hal de line Consider	so juin											149 -												11					
	1						-									-											11	1				
	1		Cotana au so Jun		160				•				-		129	1.	•						1		÷		11		-			
			Testame an 20 July for		540				1				4	50	120						10	1	T		T		1					
	200	,	Salarie de novembre	10 gh	41						1.				1								T									
	215	10	Conditation to up al	J.		. 11	2 92				1.								1.		1											
	111	11	Animoie Karlandery	r dinel			-								4. 00	4,											1					
3	191	12	ditair de dissulie	si sh	41						-		-		•		-				+.											
	345	15	d'	-		• •					-				•		1											-				-
	265	110	A uner de la C' d'amager	5/1417												-			•		-		-		•		-	-	-	-		
			Totawi an 31 x 10	1000	570		1 23				540	41	-		ma	-			1		1	-		-	T				T		T	
			0.7	1			1				T	T				-		-	1		1	-	-						10			
A CONTRACTOR OF				1																												
	1.50			-																												
Sec. 19		-		-											1	-								-								
1. A.		-		-											-	-		-						-			- 1	IF				
				1												-		-									-	IE				
			- Indiana and	Catholica											-	-																
																-								-			-	1				
				and a state																							-	I				
	1	-		1																												
																-												1				
																			-									1-	-			-

	200			1213	20172			162	- PR	To TR	Area a	1000						1		-	NO SO		- PARAMAN
															C	OMM	UNE	D	-	1	ar	nie	ver SI
2012-0224																I		É	PE		NS	-	TS
	_							ENTRET	13	-	-		-	(	nosses	REPAIL	TION	8	-	+	-		
		1	•			ENT	APPARAE DU S	na Je		TRA	TALS IN	Edott	Extra	reast or a Rabein de Cristin de	ICT 8.			TRAVAL	ts Him		Drug	tatels de	an a
	M CLEX	broates	PUBLICATION	14 13370	Gabeeler		Citta de	1 1	634.8	4-	1914			Imatal		TUTAL	*	44	NEMT			2+24 64	TOT
	100381	NUMBER	et des pièces adpresées	PROS	ana de	TRATIC Invalue rabeis	TRAVALS Both Serminols Tables	Hako	is rided		-	-	ferminfs rabais non didki	Lormin rahain ron did	in sub		rgent	Nokove	Arpet		PRATURE Bermine's Poliets Bros dedoit	termin fermin raber	a Felan
					Spat-	-	_			-		T	-	-			1	1			1	-	
			Esport des nandes natification										• •					•	• • •				
	10	4	Indian de Jamesi	3. Juni	90 .										- +					1			
	37-25	2 3	Alexandre de Gibber	nt. Chlim	90												-						
	*	4	Alani de Señas	3eMarco	.90 .		••••					-	• •			•							
		100	Erliner on 31 stress		270 .				-										• • •				
	63	5 6	delais d'aval Detations	sunt.	45 .						50		• •			1.							
	43.	7	Sulan a star	e	-15 .										• •					-			
	14	1	Hel & link Consider	so Juny		-						-7 -								-	-		
		1	tolana an so Juin		160 .					. 1	150	129-	• •			•	• •		• • •	- 1			
			betome on so dept for		<i>m</i> .							129 .				-							
	202	1	Conditation do 14 al	10 900	-11 .	111 9			1		1	• •								-	·	• •	::
	ш		Animie Karlanden	r dinel							-	4 40	,							- 1			
	345	12	dilais de diambo	1	30 .				1		1			1	• •	1				-		• •	
	263	10	Lines de la Ct d'anner	Spinie												-		-		-			
			Totany au 31 x 10		570 .	see g.	۶.,		530	1 1	51	43 90	-					-			-		
				1												+				-			
																		10		-			
	-	-																		-	IE		
				1							-									-	I		
Second Second																			-	-	1		
																				1	1		
			No. of Concession, Name			-	-	-	-	-	-	-	-							1	1	-	

日					Book1 - Exc	el			<b>b</b> –		×
File	Home	Insert Pa	age Layout	Formulas	Data	Review	View	♀ Tell me.	Javier Flores	₽ Shi	are
Paste	Calibri B I	• 11 <u>U</u> • A <sup>*</sup> <u>A</u> •	→ = = = = = = = = = = = = = = = = = = =		General \$ <b>- %</b> €.0 .00 →.0	<ul> <li>↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓</li></ul>	onditional Fo ormat as Tab ell Styles *	ormatting * le *	Helete ▼ Format ▼	Editing	
Clipboard	r <sub>a</sub>	Font	🕞 Alignr	nent 🗔	Number	r <sub>34</sub>	Styles		Cells		^
A1	* :	× ~	fx								~
A	В	С	D	E	F	G	Н	1	J	ĸ	
1											
2											
3						-	_				
4											
6		_					_				
7											
8											
9											
10											
11											
12											
13											
14											
15			124 - 52								
94 - F	She	et1 Shee	t2 Sheet3		1	:	4				F
Ready								四	-	+ 10	00%







What's wrong with the modern data stack?

The modern data stack is batch-based because streaming tools haven't exposed that same SQL interface.

# "Batch is good enough."



Replying to @sarahcat21

streams seem to me like one of those beautiful in principle and awful in practice things because of how hard it is to get right. i just can't imagine a 30-50 person company with one or two data hires running off streams but very easy to do off tables.

5:21 PM · Mar 15, 2022 · Twitter for iPhone

1 Retweet 24 Likes

### Say hello to Materialize.



### **Timely Dataflow and Differential Dataflow**

#### Timely Dataflow → "a low-latency cyclic dataflow computational model, introduced in the paper Naiad: a timely dataflow system"

### **Timely Dataflow and Differential Dataflow**

#### Timely Dataflow $\rightarrow$ computational model, introduced in the

#### Differential Dataflow $\rightarrow$

paper Naiad: a timely dataflow system"
 "a computational framework built on top of timely dataflow intended for efficiently
 performing computations on large

amounts of data and *maintaining* the computations as the data change"

"a low-latency cyclic dataflow



You can build correct, real-time applications and analytics, just using SQL + Materialize.



### Materialize's superpower: incrementally-maintained materialized views.



#### Materialized views, a primer

What's a materialized view?	A database object that contains the result of a query.
Why would I use one?	Each time you query your database you will incur a cost. Materialized views cache your results, limiting your costs and keeping your results "fresh enough."
What's the downside?	Materialized views get stale, and refreshing them can be slow and expensive.

#### Incrementally-maintained materialized views

#### What's the same?

#### What's different?

A database object that contains the result of a query.

These materialized views update their results *incrementally* as the underlying data changes. This means they are never stale, and will always return correct, up-to-date results. Even better, this means they will only perform the necessary work.

					1.175	12.986		2.66	See.	Xac	100	N.L.	Ne.s	AT A B		-								0000	RENER	North Mark
																	C	OMN	AUN	ED	1.		10	rs	nie	red \$
																		-	D	É	P	E	I	VE	SI	ES
	-			ENTRETIEN Choks								IOBSES IDBSES	REPA	RATIO	N8			+	-							
	1			DATES	-	ENT	RAPRISE I	o surra	Jas	me-	- 11	RAVAUN	EN RÉGI	IE	Extate	tabais de	19.			TRAV	AUX E	N RLCR		ENTREP	ine pe sa iatois de l	#1. N
	DE CARS	<b>DrOn.DRU</b>	INDICATION	HE L'ESTOI	Canteenie		Credat 	do	TC	Sog 3	2	() ()	4r_	-		Inavati	T	TUTA	L Adult	40_	CREW	1	-		teta de	Te
	MEROS	CT MEROS	et des pièces adrossées	des poices	ana	TRAVAC termin rabats	x TRAX	MX ints	Raboli	s rédait		-	-	- 6	ralais en déduit	terminés rabais	Nati	10	Argent	Net	Larr 1	Armi		navat's graines galais	termini cubais	ta Rolan
	N				stat-	son dre	tell son i	léðsit	Nature	Argee	1 2	Natore	Argent						1	-	T		. []_		ana did	init Nature
										1												. ,				
			Report to same and set	1										-			1200		-	-			-			
																							1			
	10	1	Unhier de Jamei	3. Jani	90 .					1:																
	23	3	Dara de firser	conficience	90									-						• •						
	16	+	Jalani de Itan	31.312000	.90 1	-			•					-	• •				•					•		
		1	Estimer on 31 detains		270 .			-																		
	63	5	Salari d'aval	same	45 ,			•							• •			1								
	\$5	6 7	Partotions	31 Jutai	45			•		1		1 50			• •			1			Ċ					
	14	1	Hel de line Conador	so juin									119													
		20% 61%		• 1								10											-1	1		
			Cotana an 30 Juin		16			*			•	4 10	129		•••					-					-	
		120	Botane an so dept fr		160 .							150	129										- 1			
	200	1	Salarie de november	80 gm	- 40 .						•	• •		•	• •				•	• •				·	• •	
	215	10	tuivoir kardandar	1- dimala									4	40									Ý			
STATISTICS.	<u>344</u>	12	Salaire de diamber	er sh	91	1.																	- 1	-		
	345	13	d"		31 .	-			• •		•	• •			•	•••		•	•		• •		-		• •	
	260	10	Funter de la C d'annadis	1 11.1			-	-			1		-	1			•	t		-	• •		-	1		1
2			Totaus au 31 x to		570 -	628 9	5.			530	1	1 50	133	40												
		200	Contract of the second						-				1			-	-					-	-	-		
ALC: NOT																										
				-		150			_				1													
2		-									100												-	-		
				-									120				1.1						-	1000		
and the C		-		1000																						
T.												-	-	-	-				1			-	-	-	-	
				-									-							-	100	-	-		1	
			- Children																							
ALCONG ALCO						-			_														1			11-

Customer	Item	Amount
Jean	Some bread	4
Marie	Most bread	7
Pierre	Bread	5
		= 16

Customer	Item	Amount
Jean	Some bread	4
Marie	Most bread	7
Pierre	Bread	5
Louise	Tiny bread	<mark>3</mark>
		= 16

How would you calculate the updated sum?

Customer	Item	Amount
Jean	Some bread	4
Marie	Most bread	7
Pierre	Bread	5
Louise	Tiny bread	<mark>3</mark>
		= <del>16 + 3</del> = 19

How would you calculate the updated sum?





### Materialize + dbt

**Because Materialize speaks** SQL, you can transform your data in real-time using Materialize + dbt.



### What's the same?

## Model definitions

dbt commands\* -

Because Materialize speaks PostgreSQL, you can define your models like you would with any other data warehouse.

Use all of the same, familiar dbt commands over streams: dbt run, dbt test, dbt docs, and more.

Documentation, lineage

Document your streaming data and get handy lineage information, as usual.

### What's different?

Materialization types

 $\rightarrow$ 

You'll want to use our custom "materializedview" materialization to create an incrementally-maintained materialized view.

### What's different

Materialization types

How some dbt commands are used

You'll want to use our custom "materializedview" materialization to create an incrementally-maintained materialized view.

Let's dig in, you're going to be pleasantly surprised!

## "dbt run" your models once and never again.



Proprietary & Confidential

dbt test works on streaming data, too!

Run dbt test on your materialized views to catch data quality issues in real time.

### dbt test + alert on streaming data #6

jwills started this conversation in Show and tell



wills 14 days ago

edited - ···

The repo for my hackday exercise is here: https://github.com/jwills/mz-hack-day-2022

Mostly copy-pasting from the Slack message where I described what I wanted to do:

"So one of the first things we do in our daily dbt run for our DWH at WeaveGrid is the standard staging work of renaming columns, cleaning up types, etc. followed by a bunch of sanity check dbt tests on that lightly processed staging data to ensure that there aren't any red flags in there that would mess up downstream table materializations. In my dream world, I would like to move that staging/testing work upstream- out of Snowflake and into dbt+materialize- so that I could run those tests "continuously" (read: with a cron job that executed like every 15 minutes-to-an-hour) to catch upstream data quality issues earlier, during business hours, and not after midnight UTC when a \*&!^?# data quality issue is going to ruin my evening.



Replying to @sarahcat21

streams seem to me like one of those beautiful in principle and awful in practice things because of how hard it is to get right. i just can't imagine a 30-50 person company with one or two data hires running off streams but very easy to do off tables.

5:21 PM · Mar 15, 2022 · Twitter for iPhone

1 Retweet 24 Likes



#### Jessica Laughlin @JLDLaughlin · Mar 15

you would be able to if the streaming system was hosted and providing the right interface - SQL! totally doable with a single dev.

6





2

....

ıll

...

Replying to @JLDLaughlin and @sarahcat21

one day i will try this and if you live up to your promise i don't know what i will do but it'll be something

7:56 PM · Mar 15, 2022 · Twitter for iPhone

## Thank you!

Questions? Come to our booth or find me at @JLDLaughlin or jessica@materialize.com

We're hiring! jobs@materialize.com

